



Hi-C detects genomic structural variants in peripheral blood of pediatric leukemia patients

Claire Mallard,^{1,2,7} Michael J. Johnston,^{1,2,7} Anna Bobyn,^{1,2} Ana Nikolic,^{1,2,3} Bob Argiropoulos,^{2,4} Jennifer A. Chan,^{1,2,5} Gregory M.T. Guilcher,^{1,2,6} and Marco Gallo^{1,2,3}

¹Arnie Charbonneau Cancer Institute, ²Alberta Children's Hospital Research Institute, ³Department of Molecular Biology and Biochemistry, ⁴Department of Medical Genetics, ⁵Department of Oncology, ⁶Department of Pediatrics, Cumming School of Medicine, University of Calgary, Calgary, Alberta T2N 1N4, Canada

Abstract B-cell acute lymphoblastic leukemia (B-ALL) is often driven by chromosome translocations that result in recurrent and well-studied gene fusions. Currently, fluorescent in situ hybridization probes are used to detect candidate translocations in bone marrow samples from B-ALL patients. Recently Hi-C, a sequencing-based technique originally designed to reconstruct the three-dimensional architecture of the nuclear genome, was shown to effectively recognize structural variants. Here, we demonstrate that Hi-C can be used as a genome-wide assay to detect translocations and other structural variants of potential clinical interest. Structural variants were identified in both bone marrow and peripheral blood samples, including an *ETV6–RUNX1* translocation present in one pediatric B-ALL patient. Our report provides proof of principle that Hi-C could be an effective strategy to globally detect driver structural variants in B-ALL peripheral blood specimens, reducing the need for invasive bone marrow biopsies and candidate-based clinical tests.

Corresponding authors:
marco.gallo@ucalgary.ca;
greg.guilcher@
albertahealthservices.ca

[Supplemental material is available for this article.]

© 2022 Mallard et al. This article is distributed under the terms of the Creative Commons Attribution-NonCommercial License, which permits reuse and redistribution, except for commercial purposes, provided that the original author and source are credited.

Ontology term: hematological neoplasm

Published by Cold Spring Harbor Laboratory Press

doi:10.1101/mcs.a006157

INTRODUCTION

B-cell acute lymphoblastic leukemia (B-ALL) is a malignancy of the bone marrow and blood that results in the uncontrolled proliferation of B-cell lineage progenitors. The frequency distribution of B-ALL diagnoses is bimodal, with peaks in childhood and in adults ~50 yr of age (for review, see Terwilliger and Abdul-Hay 2017). New treatment approaches have considerably improved overall outcomes in children, with 98% of pediatric patients going into remission following treatment and a cure rate of 90%. On the other hand, clinical management of adult B-ALL has proven more challenging, with median overall survival of 11 mo.

At the genomic level, B-ALL is primarily driven by translocations and other structural variants (SVs) that are thought to originate in B-lymphocyte progenitor cells. Some recurrent translocations occur with high frequency in the patient population and can result in tumorigenic, gain-of-function fusion proteins. These genomic events include t(9;22) (a translocation

⁷Co-first authors

between Chromosome 9 and Chromosome 22, also known as the Philadelphia chromosome) resulting in *BCR-ABL1* fusion; t(12;21), resulting in the *ETV6-RUNX1* fusion; t(1;19), resulting in the *TCF3-PBX1* fusion; and rearrangements of the *MLL* gene with many fusion partners (for review, see Woo et al. 2014). Detection of specific translocation(s) and SVs present at diagnosis allows molecular subtyping of the B-ALL case, which informs patient prognosis and guides therapy selection.

Clinical tests for the detection of translocations mostly rely on fluorescent in situ hybridization (FISH) of bone marrow samples. Bone marrow cells are obtained via pelvic bone tap, which is an invasive and painful procedure. Because this is a candidate-based clinical test, multiple FISH probes must be used to assay for candidate translocations. Karyotyping is also performed on these samples, but this method may miss SVs that are not big enough to be detected. Being able to capture different types of SVs—including smaller events—in one individual test could improve the speed and precision of diagnosis.

In this report, we test the potential of Hi-C (Belton et al. 2012)—a chromosome conformation capture technique—to provide effective genome-wide information on the SV status of B-ALL samples. Hi-C was originally designed to assess physical interactions between non-contiguous chromosome regions (Dixon et al. 2012; Rao et al. 2014). This information is then used to computationally reconstruct three-dimensional (3D) genome architecture (i.e., how DNA folds within the nucleus). In essence, Hi-C involves cross-linking DNA within cells to create a snapshot of which genomic regions physically interact with each other. When libraries of interacting regions are sequenced very deeply, it is possible to reconstruct both large-scale genomic structures (domains and compartments) and finer features like chromatin loops formed by *cis*-regulatory regions interacting with their target promoters (Rao et al. 2014; Lu et al. 2020).

Hi-C can visualize major SVs in a sample (Dixon et al. 2018). For instance, translocations would appear as interactions between regions normally situated on different chromosomes with a characteristic butterfly appearance on a Hi-C contact map (Fig. 1). Recent publications have shown that Hi-C enables detection of SVs even with relatively shallow sequencing of the libraries (Harewood et al. 2017; Dixon et al. 2018).

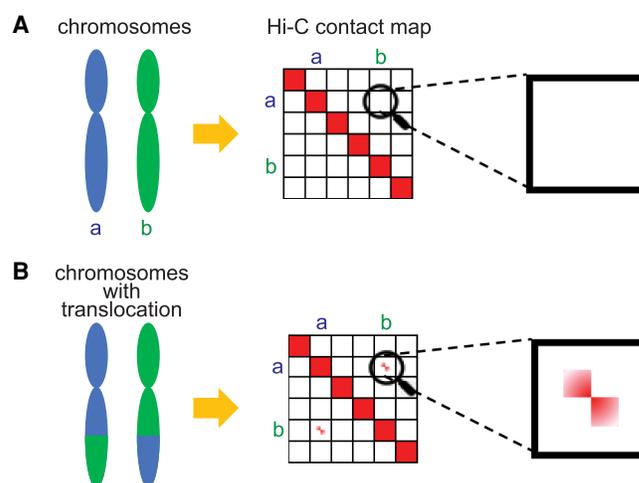


Figure 1. Identification of chromosomal translocations with Hi-C. (A) Diagram illustrating the expected Hi-C contact frequency maps in the case of a genome with no translocations. A normal karyotype will result in minimal interchromosomal contacts. (B) Diagram illustrating the expected Hi-C contact frequency maps in the case of a genome with a balanced translocation. Elevated interchromosomal contacts would be observed. A balanced translocation would appear as a contact with a characteristic “butterfly” appearance (*inset*).

Table 1. Patient and sample metadata

Patient ID	Diagnosis	Type of sample	Sex
4437	B-ALL	Marrow and peripheral blood	Female
4439	B-ALL	Peripheral blood	Male
4440	B-ALL	Peripheral blood	Male
4441	B-ALL	Peripheral blood	Female

Here, we assess the feasibility of using shallow (low-coverage) Hi-C to infer translocations and other SVs in a discovery cohort of children with B-ALL. We also test the feasibility of detecting SVs by performing Hi-C on peripheral blood of B-ALL patients as a less-invasive alternative to using bone marrow aspirates.

RESULTS

Hi-C Data Sets of Bone Marrow and Peripheral Blood Samples from Pediatric B-ALL Patients

We generated Hi-C libraries for five samples from four pediatric B-ALL patients (Table 1) with an average contact resolution of ~35 kb. For all the relevant statistics related to the Hi-C libraries, please see Supplemental Table S1. For patient 4437, we generated Hi-C data on both bone marrow cells (4437M) and peripheral blood cells (4437B) to test our hypothesis that translocations and structural variants can be detected in leukemic blasts that circulate in the peripheral blood. For all the other patients, we generated Hi-C data sets from their peripheral blood. For each patient, we also had clinical data about their B-ALL diagnosis using FISH (Table 2). Our cohort included two males and two females. Samples from three patients were collected at the initial diagnosis of leukemia. The sample from patient 4441 was obtained when that individual was in remission.

Hi-C maps generated from all samples were informative of the overall genomic structure of all samples (Fig. 2A–E). They all show strong interactions primarily within each chromosome, as expected, including stronger interactions that preferentially occur within the p-arm and q-arm of each chromosome. However, it is also possible to notice putative interactions between chromosomes, which likely represent SVs. A candidate SV, for instance, appears as a translocation between the q-arm of Chromosome 1 and Chromosome 9 in patient 4440 (Fig. 2D).

Table 2. Comparison of structural variants detected by conventional cytogenetic analysis and Hi-C

Sample ID	Structural variants (SVs)	Detected with clinical test	Detected by Hi-C
4437M	t(12:21) (<i>ETV6:RUNX1</i>)	Yes	Yes
	Intrachromosomal SVs on Chr 12 near <i>ETV6</i>	No	Yes
4437B	t(12:21) (<i>ETV6:RUNX1</i>)	Yes	Yes
4439	Intrachromosomal SVs on Chr 21q near <i>RUNX1</i>	Unexplained signal	Explained signal
	Deletion at 21q22.3	No	Yes
4440	Intrachromosomal SVs on Chr 14q12	No	Yes
4441	NA		

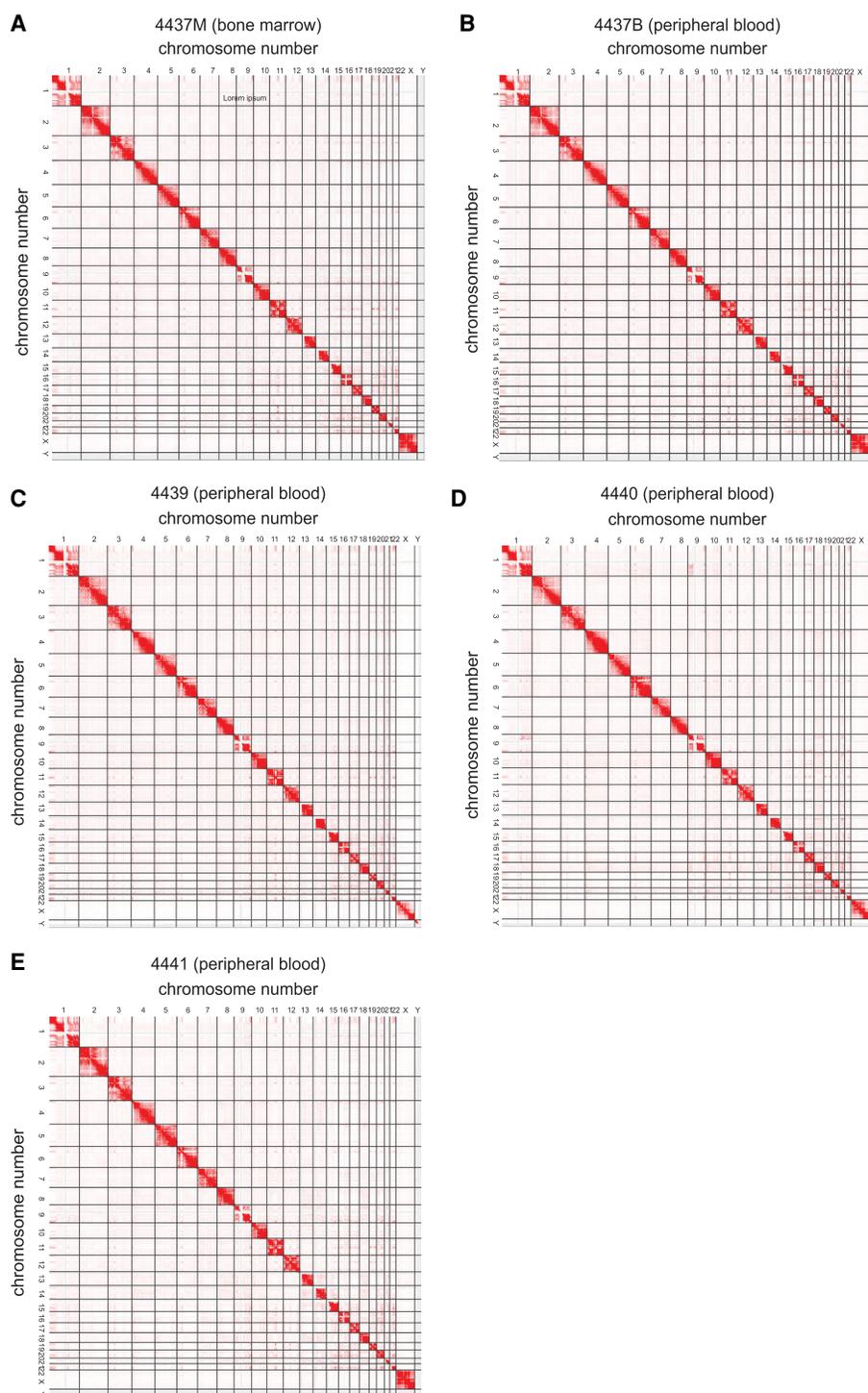


Figure 2. Hi-C contact maps for five samples collected from pediatric B-ALL patients. (A) Contact map for Hi-C data generated from the bone marrow sample of patient 4437. (B) Contact map for Hi-C data generated from the peripheral blood sample of patient 4437. (C) Contact map for Hi-C data generated from the peripheral blood sample of patient 4439. (D) Contact map for Hi-C data generated from the peripheral blood sample of patient 4440. (E) Contact map for Hi-C data generated from the peripheral blood sample of patient 4441.

We uploaded the processed Hi-C maps for all five samples to the WashU Epigenome Browser (https://wangftp.wustl.edu/hubs/gallo_B-ALL) to facilitate public access to our data sets and their downstream analyses by users. The raw Hi-C data are also available (see Data Deposition and Access section). We hope these data sets will be a resource for the community.

Identification of Translocations Using Hi-C Data Sets

As a first test of our data sets, we wanted to determine if B-ALL-related translocations could be detected using Hi-C data from the bone marrow sample we profiled. Clinical FISH probes confirmed that leukemic blasts from patient 4437 were positive for the *ETV6*–*RUNX1* translocation, which is a driver mutation for B-ALL and occurs via a balanced exchange of DNA between Chromosomes 12 and 21. Focusing on these chromosomes, we were able to identify this translocation, which appeared as a butterfly-shaped signal, as expected (Fig. 3A). Furthermore, our data sets allowed the identification of the precise breakpoints of the *ETV6*–*RUNX1* translocation in this patient. The Hi-C contact matrix showed breakpoints corresponding to intron 5 of *ETV6* and intron 2 of *RUNX1* (Fig. 3B). These results support previous findings on this well-researched translocation.

Next, for the same patient, we looked for that same translocation using Hi-C data from peripheral blood. Although peripheral blood has a lower density of leukemic blasts than the bone marrow, we were still able to identify a significant signal at the expected site of translocation (Fig. 3C). This finding is further supported by comparing the signals observed in the translocation-positive patient (4437) and a translocation-negative patient (4441). At the *ETV6*–*RUNX1* intersection locus for patient 4441, there is no significant signal from the peripheral blood Hi-C data (Fig. 3D). Altogether, our results support the notion that shallow Hi-C has sufficient sensitivity to detect SVs, even in peripheral blood samples from B-ALL patients.

Identification of Complex SVs in Pediatric B-ALL Samples Using Hi-C

Having established that shallow Hi-C enables the detection of known B-ALL translocations, we then wondered if we could use our data sets to identify other SVs in the genomes of our patients. Examining our data sets, we were able to identify SVs that were not reported following clinical tests for the same patients (Table 2). These SVs appeared as strong signals off the diagonal of Hi-C contact matrices. Among the signals identified were candidate tandem duplications, inversions, and deletions, but further analyses would be required to fully characterize each SV. However, the detection of abnormalities and the genes involved at these loci are a starting point for further analysis into these cancer genotypes that would not be possible to explore using only clinical FISH probes.

For patient 4439, complex intrachromosomal structural variants were detected around the *RUNX1* locus on Chromosome 21 (Fig. 4A). Some cytogenic abnormalities were detected near this locus in clinical analysis, but Hi-C was able to specifically identify the breakpoints. It is possible that this SV could affect *RUNX1* expression, but we cannot confirm this hypothesis with the data we currently have. Further, the effects seen line up closely with a known B-ALL cytogenetic subgroup known as iAMP21—intrachromosomal amplification of Chromosome 21 (Harrison 2015). Further, this subgroup has a threefold higher relapse rate and lower 5-yr survival (71%) than other B-ALL patients (Harrison 2015), making the rapid diagnosis and characterization of leukemia critical for proper diagnosis and clinical management of these high-risk patients.

Further structural variants were detected in the bone marrow Hi-C contact matrix for sample 4437M, in which we previously detected the *ETV6*–*RUNX1* translocation (Fig. 4B). These abnormalities were detected as discontinuous, long-range intrachromosomal interactions

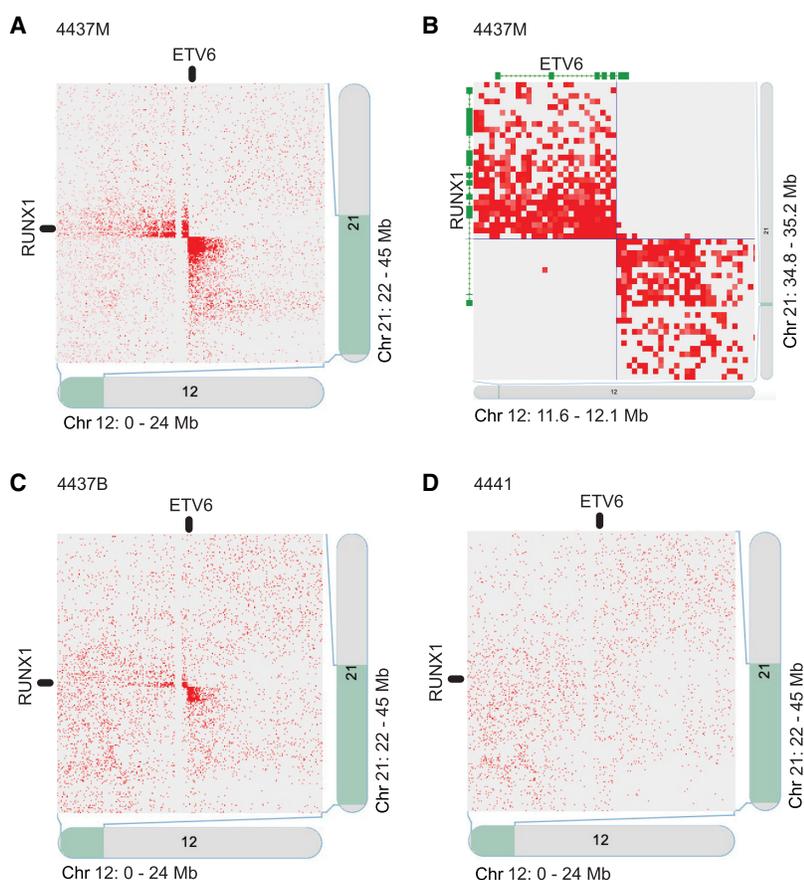


Figure 3. The *ETV6*–*RUNX1* translocation was detected in the bone marrow and peripheral blood Hi-C contact maps. (A) The interchromosomal contact map shows the interactions between Chr 12 (0–24 Mb) and Chr 21 (22–45 Mb) for the bone marrow sample of patient 4437. The bright red butterfly shape shows the balanced translocation at the *ETV6* locus (Chr 12) and the *RUNX1* locus (Chr 21). (B) Higher-magnification view of the *RUNX1*–*ETV6* translocation detected by Hi-C in the bone marrow sample of patient 4437 (previous panel). The *ETV6* and *RUNX1* genes are shown to illustrate how the data can clearly identify the breakpoint of the translocation. The breakpoint for *ETV6* occurs in intron 5 and the breakpoint for *RUNX1* occurs in intron 2. (C) Interchromosomal contact map showing interactions between Chr 12 (0–24 Mb) and Chr 21 (22–45 Mb) for the peripheral blood sample of patient 4437. The bright red butterfly shape shows the balanced translocation at the *ETV6* locus (Chr 12) and the *RUNX1* locus (Chr 21). (D) Interchromosomal contact map for Chr 12 (0–24 Mb) and Chr 21 (22–45 Mb) for the peripheral blood sample of patient 4441. The absence of abnormal signal signifies that no translocation between these loci is present in this sample.

along Chromosome 12 near the *ETV6* locus, suggesting that the local chromatin environment around *ETV6* may be further disrupted independent of the *ETV6*–*RUNX1* fusion. Further, the breakpoints of another potential SV (lower interaction box in Fig. 4B) correspond to the genes *APOBEC1* and *LMO3*. Both genes have been identified as important for leukemias in previous studies. *APOBEC1* encodes a cytosine deaminase and has been linked to cancer development in acute myeloid leukemia (Saraconi et al. 2014). *LMO3* codes for a transcription factor and oncogene that contributes to the etiology of T-cell acute lymphoblastic leukemia (La Starza et al. 2020).

We also identified a complex SV on the p-arm of Chromosome 14 of the sample from patient 4440 (Fig. 4C). This cluster of abnormal Hi-C signals is consistent with a chromothripsis event, in which the chromosome was partially shattered and many gene loci were altered

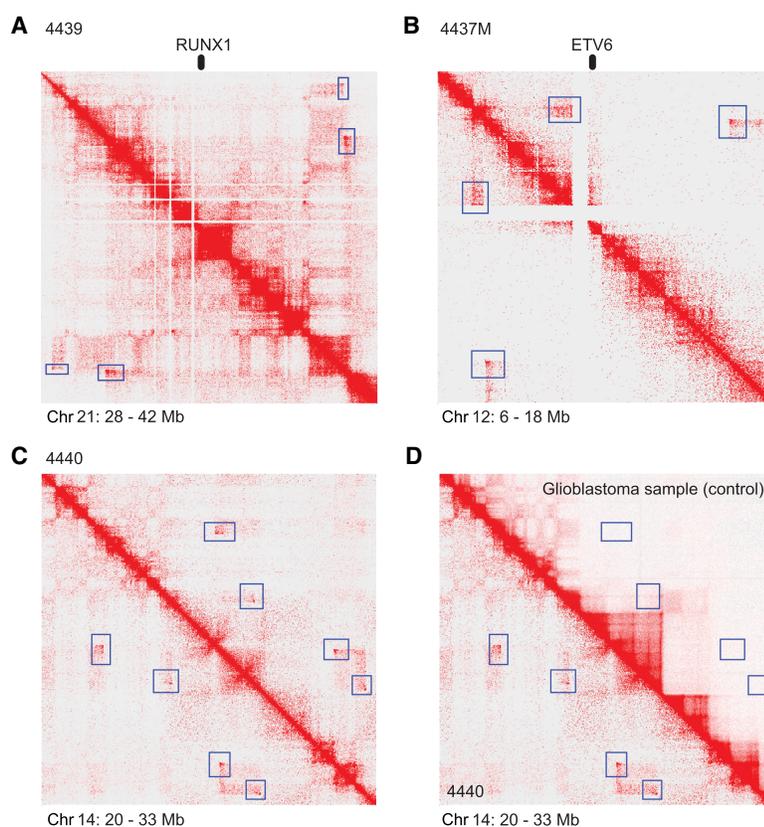


Figure 4. Hi-C enables genome-wide identification of structural variants. (A) Intrachromosomal structural variants near the *RUNX1* locus on Chromosome 21 (28–42 MB) for patient 4439. (B) Intrachromosomal structural variants near the *ETV6* locus on Chromosome 12 (6–18 MB) in the bone marrow sample of patient 4437. (C) Intrachromosomal structural variants on Chromosome 14 (20–33 MB) for patient 4440. (D) Intrachromosomal structural variants on Chromosome 14 (20–33 MB) for patient 4440 (*bottom left*) compared with the same region on a GBM patient Hi-C contact map (*top right*). For all panels, the putative structural variants are highlighted in blue boxes.

all at once. This SV was not observed in previously published Hi-C data from a glioblastoma sample (G523) (Johnston et al. 2019), which has a nonaltered chromosomal structure in this region (upper half of Fig. 4D).

3D Modeling of B-ALL Chromatin Interactions

In addition to large-scale SVs, Hi-C provides information on chromatin interactions along chromosome regions that do not harbor obvious SVs (Dixon et al. 2012; Rao et al. 2014). 3D chromatin interactions result in the formation of domains and loops, which are important factors for gene expression and determination of cell states (Dixon et al. 2015). We decided to test whether our shallow Hi-C data were sufficient to probe chromatin interactions in B-ALL patient samples. To achieve this goal, we used CSynth (Todd et al. 2021), a computational tool that generates 3D models of chromosomes based on data generated from chromosomal conformation techniques like Hi-C. We modeled a region of Chromosome 12 (77–86 MB) and looked for differences in loop structure between bone marrow cells (Fig. 5A) and peripheral blood cells (Fig. 5B) for patient 4437. We expected to see differing loop structures between the two because 3D chromatin looping structures are cell type-specific. We

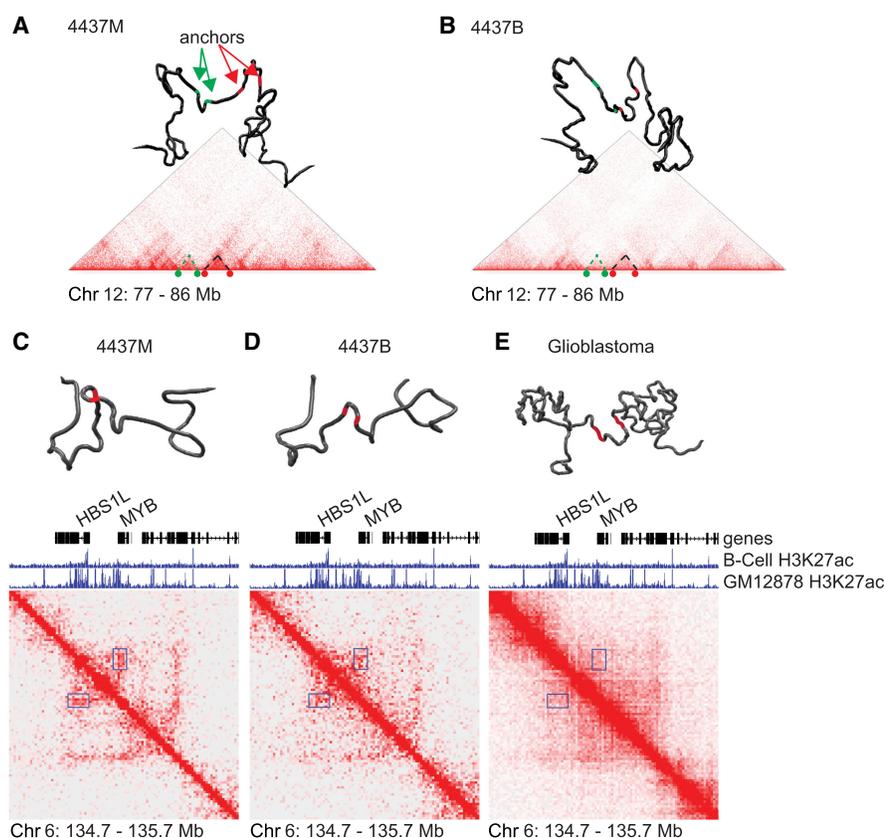


Figure 5. 3D modeling of shallow Hi-C data reveals information on gene regulation. (A,B) Hi-C-based 3D models of Chromosome 12 (77–86 Mb) from patient 4437 displaying both bone marrow (A) or peripheral blood (B). Green highlights the loop boundaries for the loop spanning 77.1–77.8 MB, and red highlights the loop boundaries for the loop spanning 78.1–79.4 MB. (C–E) Hi-C contact matrices and 3D models of Chromosome 6: 134.7–135.7 MB for patient 4437 bone marrow (C), patient 4437 peripheral blood (D), and a control glioblastoma patient (E). The loop formed between an H3K27ac enhancer region and the gene *MYB* is highlighted in the blue box on the Hi-C matrix and by red loop boundaries in the 3D model. Chromatin immunoprecipitation with sequencing (ChIP-seq) tracks for the histone mark H3K27ac for a primary B-cell sample and a cell line (GM12878) are shown in each panel.

highlighted two pairs of loop anchors corresponding to two loops: 77.1–77.8 MB and 78.1–79.4 MB (Fig. 5A,B; green and red, respectively). In Hi-C, the intensity of the loops was higher in the bone marrow data, likely because of the different densities of bone marrow-specific cells, such as B-lymphocytes. This was supported by the 3D models, which showed tighter and more distinct loops in the bone marrow model than in the peripheral blood model, as seen by the closer proximity of the green and red loop boundaries.

Next, we wanted to see if this phenomenon would be apparent for specific enhancer to gene loops related to leukemias. We identified putative *cis*-regulatory regions using chromatin immunoprecipitation with sequencing (ChIP-seq) data for B cells and GM12878, a lymphoblastoid cell line, generated by ENCODE (Davis et al. 2018). Specifically, we looked at the acetylation of histone 3 on lysine 27 (H3K27ac), which marks enhancer and promoter regions. We found a strong loop present from the bone marrow Hi-C between a putative enhancer marked by H3K27ac and the proto-oncogene *MYB*, which is overexpressed in leukemias (Lahortiga et al. 2007). In the Hi-C contact matrix for patient 4437 bone marrow, there was a strong signal at the interaction locus for the enhancer and *MYB* (Fig. 5C). This

corresponded to an obvious loop structure in the 3D model (Fig. 5C). In the peripheral blood Hi-C contact matrix for the same patient, there was a signal present but with lower intensity. The 3D model also showed the loop, but the boundaries were farther apart (Fig. 5D). Finally, this loop was not present in the Hi-C contact matrix of a previously published glioblastoma sample (Fig. 5E). These results demonstrate that shallow Hi-C data sets of clinical samples are sufficient to reconstruct the 3D milieu of specific genomic regions that are relevant to disease etiology.

DISCUSSION

Hi-C has been shown to be an effective genomic approach to solve the 3D architecture of the genome and also to identify large SVs (Harewood et al. 2017; Dixon et al. 2018). We reasoned that the unbiased, genome-wide assessment of structural variation offered by Hi-C could be an asset in the diagnosis and molecular subtyping of malignancies driven by defined SVs, especially translocations. We, therefore, decided to focus our efforts on B-ALL, which is very often driven by translocation events that generate gene fusions that maintain malignant cells in progenitor states.

Balanced translocations can easily be visualized in Hi-C contact maps by virtue of their characteristic “butterfly” shape. One of our patients had a canonical translocation between Chromosomes 12 and 21, resulting in the *ETV6–RUNX1* gene fusion. This translocation was confirmed by FISH. Our shallow Hi-C approach was able to clearly identify this translocation in the bone marrow sample by simple visual inspection. Importantly, we were also able to detect this translocation in the peripheral blood sample from the same patient. Additionally, we successfully identified other putative SVs in the peripheral blood samples of the other two patients. The blood sample from an additional patient was collected when the child was in remission and, as expected, did not have obvious abnormal Hi-C signals. Overall, our data support the suitability of shallow Hi-C as a tool to detect known and unknown SVs in B-ALL genomes.

Although we only had one patient-matched pair of bone marrow and peripheral blood specimens, our results provide proof of principle that shallow Hi-C can be successfully applied to peripheral blood samples to detect recurrent translocations in B-ALL patients. The success of this approach depends on the burden of leukemic blasts in circulation and could correlate with the disease stage. We expect that simple refinements of the approach we describe could significantly improve the power to detect translocations using peripheral blood. For instance, a simple solution would be to enrich leukemic blasts by performing fluorescence-activated cell sorting directly on peripheral blood (Rezaei et al. 2003; Cheng et al. 2019). This approach would reduce the need for bone marrow biopsies, which can be painful procedures. Overall, Hi-C could be a useful method for the detection of SVs in samples with high leukemic blast representation, whereas FISH provides single-cell resolution and could therefore be more useful in cases with low blast counts—for instance, in relapse or refractory settings.

In patients who did not have the recurrent *ETV6–RUNX1* translocation, we also detected putative SVs near *ETV6* and *RUNX1*. This finding is interesting because it suggests that there may be alternative mechanisms of disrupting the proper expression of these genes in addition to the canonical translocation. Hi-C allows the visualization of large genomic abnormalities and could provide information on the potential disruption of transcriptional programs associated with these genes. Leukemias that converge on disruption of transcriptional pathways mediated by *ETV6* and *RUNX1*, whether through translocations or other means, could be identified.

Shallow Hi-C does have some limitations. These include the specialized informatic analysis required and its inability to detect very small SVs (e.g., single-gene duplications or very small deletions). Hi-C output is also relatively qualitative and visual inspection is required for inferences of genomic structural anomalies. However, future studies could lead to standard computational pipelines for the accurate identification of SVs using Hi-C. Although a powerful technique to detect structural variants like translocations and inversions, Hi-C is not optimal to identify copy-number changes, including heterozygous deletions and amplifications. This is especially true for the shallow Hi-C sequencing approach we took in our study. Therefore, this method may not be applicable to B-ALL cases that do not harbor translocations but that instead present with hyperdiploid or hypodiploid genomes and/or more subtle copy-number gains and losses.

Depending on the local health-care model and cost structure, shallow Hi-C could be cost-competitive with other sequencing-based methods that are currently in use for diagnosis and molecular subtyping of leukemias (Inaba et al. 2017). The cost for Hi-C library preparation and sequencing was ~\$2000 per sample for the cases described in this study at 35-kb resolution. This cost could be significantly reduced by multiplexing samples on sequencing runs and reducing the targeted Hi-C map resolution. As an example, the *ETV6*–*RUNX1* translocation detected in one of our cases was clearly detectable at a resolution of 250 kb. Although shallow whole-genome sequencing provides genome-wide information on SV status, Hi-C provides a more straightforward platform to reconstruct complex structural events (Dixon et al. 2018). We also tried to compare the cost of shallow Hi-C to FISH. Based on our experience, clinical evaluation of FISH probes costs ~\$450 per candidate SV. Typically, multiple candidate loci are assayed for each patient (e.g., t(9;22), t(4;11), and t(12;21)) with each candidate acting as a multiplier on this cost. Overall, the cost of performing multiple FISH experiments is comparable to that of shallow Hi-C, whereas only the latter provides an unbiased assessment of genomic events at noncandidate loci.

Additional studies with larger cohorts of patients will be required, but our results suggest that shallow Hi-C could be used to reinforce the standard of care testing (flow cytometry and cytogenetics) and hopefully implemented in the clinic to test for translocations using peripheral blood of B-ALL patients. The standardization of Hi-C library preparation and its related computational pipelines, and the wide availability of sequencing in clinical settings, make this option worth exploring.

METHODS

Generation of Hi-C Data Sets

Peripheral or marrow blood samples were collected from patients with B-ALL at the Alberta Children's Hospital. Blood samples were diluted 1:1 with phosphate-buffered saline (PBS) + 2% fetal bovine serum (FBS). Peripheral blood mononuclear cells were isolated using SepMate columns (Stemcell Technologies). Cells were rinsed and resuspended to 25 mL in PBS + 2% FBS. Cell density was measured using a Countess cell counter (ThermoFisher). One to five million cells were then used as input for Hi-C library construction using the Arima Hi-C Kit (Arima Genomics), including KAPA Hyper Prep indexing and library amplification.

Samples were sequenced 2 × 150 bp on a NextSeq 500 instrument (Illumina) at the Center for Health Genomics and Informatics (CHGI) at the University of Calgary. Reads were processed using Juicer (Durand et al. 2016b), aligning to reference genome hg38 and the restriction enzyme DpnII, to generate Hi-C contact matrices (.hic files). Contact matrices were visualized using Juicebox (Durand et al. 2016a).

SV Detection

We considered that genomic structural variants would appear as long-range interactions on a Hi-C contact map. Using Juicebox, we visually inspected the whole-genome contact map using balanced normalization at 100-kbp resolution for abnormally high interchromosomal signals located >5 Mbp apart. A resolution of 50-kbp was used to inspect abnormal intra-chromosomal signals. Additionally, putative SV signals were inspected to ensure that the Observed/Expected ratios were much greater than 1. Finally, putative SV signals were compared to negative control Hi-C data sets in which leukemic disease had cleared (sample 4441) or which would not be expected to share leukemic genomic alterations (published glioblastoma data set from sample G523; Johnston et al. 2019).

3D Modeling of Hi-C Data

Hi-C contact matrices and BED files containing loop loci of interest were loaded into CSynth (Todd et al. 2021) (csynth.org). Default parameters were used with the following exceptions: push apart force: 3×10^{-4} ; contact force: 52; diameter: 20.

H3K27ac ChIP-seq Data sets

We accessed H3K27ac fold enrichment traces from ENCODE (Davis et al. 2018) through Juicebox with the following identifiers: ENCFF696PMM (*Homo sapiens* B-cell female adult (27 yr of age); <https://www.encodeproject.org/files/ENCFF696PMM/>); ENCFF340JIF (*Homo sapiens* GM12878; <https://www.encodeproject.org/files/ENCFF340JIF/>).

ADDITIONAL INFORMATION

Data Deposition and Access

All Hi-C data sets generated for this study are available to the community. The processed Hi-C data can be loaded into the WashU Epigenome Browser from the following link: https://wangftp.wustl.edu/hubs/gallo_B-ALL. The raw Hi-C data have been uploaded at the EGA (<https://ega-archive.org/>) repository (accession number EGAS00001005605).

Ethics Statement

Collection and experimental use of all samples were approved by the Health Research Ethics Board of Alberta—Cancer Committee (certificate number HREBA.CC-18-0169). All methods were performed according to the relevant guidelines and regulations as they pertain to the generation and storage of genomic data sets at the University of Calgary. All samples were collected after informed consent was obtained from each participant or their legal representatives.

Author Contributions

M.G. and G.M.T.G. conceived the study. G.M.T.G. and J.A.C. collected and banked the patient specimens. C.M., M.J.J., A.B., and B.A. performed the experimental activities. M.G., C.M., and M.J.J. wrote the manuscript. All coauthors contributed to editing the manuscript.

Funding

This project was funded by a Young Investigator Award in Cell and Gene Therapy for Cancer grant from the Alliance for Cancer Gene Therapy (ACGT) to M.G.; a Discovery Grant from the Natural Science and Engineering Research Council (NSERC) to M.G.; a Canada Research Chair from the Government of Canada to M.G.; a USRA studentship from NSERC to C.M.; a postdoctoral fellowship from the Canadian Institutes of Health Research (CIHR) to M.J.J.; and support from the Alberta Children's Hospital Foundation and the Kids Cancer Care Foundation of Alberta to J.A.C. for the Brain Tumor and Tissue Bank.

Competing Interest Statement

The authors have declared no competing interest.

Received October 22, 2021;
accepted in revised form
November 22, 2021.

REFERENCES

- Belton J-M, McCord RP, Gibcus JH, Naumova N, Zhan Y, Dekker J. 2012. Hi-C: a comprehensive technique to capture the conformation of genomes. *Methods* **58**: 268–276. doi:10.1016/j.ymeth.2012.05.001
- Cheng J, Klairmont MM, Choi JK. 2019. Peripheral blood flow cytometry for the diagnosis of pediatric acute leukemia: highly reliable with rare exceptions. *Pediatr Blood Cancer* **66**: e27453. doi:10.1002/pbc.27453
- Davis CA, Hitz BC, Sloan CA, Chan ET, Davidson JM, Gabdank I, Hilton JA, Jain K, Baymuradov UK, Narayanan AK, et al. 2018. The Encyclopedia of DNA elements (ENCODE): data portal update. *Nucl Acids Res* **46**: D794–D801. doi:10.1093/nar/gkx1081
- Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y, Hu M, Liu JS, Ren B. 2012. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* **485**: 376–380. doi:10.1038/nature11082
- Dixon JR, Jung I, Selvaraj S, Shen Y, Antosiewicz-Bourget JE, Lee AY, Ye Z, Kim A, Rajagopal N, Xie W, et al. 2015. Chromatin architecture reorganization during stem cell differentiation. *Nature* **518**: 331–336. doi:10.1038/nature14222
- Dixon JR, Xu J, Dileep V, Zhan Y, Song F, Le VT, Yardimci GG, Chakraborty A, Bann D V, Wang Y, et al. 2018. Integrative detection and analysis of structural variation in cancer genomes. *Nat Genet* **50**: 1388–1398. doi:10.1038/s41588-018-0195-8
- Durand NC, Robinson JT, Shamim MS, Machol I, Mesirov JP, Lander ES, Aiden EL. 2016a. Juicebox provides a visualization system for Hi-C contact maps with unlimited zoom. *Cell Syst* **3**: 99–101. doi:10.1016/j.cels.2015.07.012
- Durand NC, Shamim MS, Machol I, Rao SSP, Huntley MH, Lander ES, Aiden EL. 2016b. Juicer provides a one-click system for analyzing loop-resolution Hi-C experiments. *Cell Syst* **3**: 95–98. doi:10.1016/j.cels.2016.07.002
- Harewood L, Kishore K, Eldridge MD, Wingett S, Pearson D, Schoenfelder S, Collins VP, Fraser P. 2017. Hi-C as a tool for precise detection and characterisation of chromosomal rearrangements and copy number variation in human tumours. *Genome Biol* **18**: 125. doi:10.1186/s13059-017-1253-8
- Harrison CJ. 2015. Blood spotlight on iAMP21 acute lymphoblastic leukemia (ALL), a high-risk pediatric disease. *Blood* **125**: 1383–1386. doi:10.1182/blood-2014-08-569228
- Inaba H, Azzato EM, Mullighan CG. 2017. Integration of next-generation sequencing to treat acute lymphoblastic leukemia with targetable lesions: the St. Jude Children’s Research Hospital approach. *Front Pediatr* **5**: 258. doi:10.3389/fped.2017.0025
- Johnston MJ, Nikolic A, Ninkovic N, Guilhamon P, Cavalli FMG, Seaman S, Zemp FJ, Lee J, Abdelkareem A, Ellestad K, et al. 2019. High-resolution structural genomics reveals new therapeutic vulnerabilities in glioblastoma. *Genome Res* **29**: 1211–1222. doi:10.1101/gr.246520.118
- La Starza R, Pierini V, Pierini T, Nofrini V, Matteucci C, Arniani S, Moretti M, Lema Fernandez AG, Pellanera F, Di Giacomo D, et al. 2020. Design of a comprehensive fluorescence in situ hybridization assay for genetic classification of T-cell acute lymphoblastic leukemia. *J Mol Diagn* **22**: 629–639. doi:10.1016/j.jmoldx.2020.02.004
- Lahortiga I, De Keersmaecker K, Van Vlierberghe P, Graux C, Cauwelier B, Lambert F, Mentens N, Beverloo HB, Pieters R, Speleman F, et al. 2007. Duplication of the MYB oncogene in T cell acute lymphoblastic leukemia. *Nat Genet* **39**: 593–595. doi:10.1038/ng2025
- Lu L, Liu X, Huang W-K, Giusti-Rodríguez P, Cui J, Zhang S, Xu W, Wen Z, Ma S, Rosen JD, et al. 2020. Robust Hi-C maps of enhancer-promoter interactions reveal the function of non-coding genome in neural development and diseases. *Mol Cell* **79**: 521–534.e15. doi:10.1016/j.molcel.2020.06.007
- Rao SSP, Huntley MH, Durand NC, Stamenova EK, Bochkov ID, Robinson JT, Sanborn AL, Machol I, Omer AD, Lander ES, et al. 2014. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* **159**: 1665–1680. doi:10.1016/j.cell.2014.11.021
- Rezaei A, Adib M, Mokarian F, Tebianian M, Nassiri R. 2003. Leukemia markers expression of peripheral blood vs bone marrow blasts using flow cytometry. *Med Sci Monit* **9**: CR359-62.
- Saraconi G, Severi F, Sala C, Mattiuz G, Conticello SG. 2014. The RNA editing enzyme APOBEC1 induces somatic mutations and a compatible mutational signature is present in esophageal adenocarcinomas. *Genome Biol* **15**: 417. doi:10.1186/s13059-014-0417-z
- Terwilliger T, Abdul-Hay M. 2017. Acute lymphoblastic leukemia: a comprehensive review and 2017 update. *Blood Cancer J* **7**: e577. doi:10.1038/bcj.2017.53
- Todd S, Todd P, McGowan SJ, Hughes JR, Kakui Y, Leymarie FF, Latham W, Taylor S. 2021. CSynth: an interactive modelling and visualization tool for 3D chromatin structure. *Bioinformatics* **37**: 951–955. doi:10.1093/bioinformatics/btaa757
- Woo JS, Alberti MO, Tirado CA. 2014. Childhood B-acute lymphoblastic leukemia: a genetic update. *Exp Hematol Oncol* **3**: 16. doi:10.1186/2162-3619-3-16